Biophysical Journal Volume 90 March 2006 1895–1902

Force Unfolding Single RNAs

Fei Liu,* Huan Tong,* and Zhong-can Ou-Yang*†
*Center for Advanced Study, Tsinghua University, Beijing 100084, China; and †Institute of Theoretical Physics, The Chinese Academy of Sciences, Beijing 100080, China

ABSTRACT We develop a continue time Monte Carlo algorithm to simulate single RNAs unfolded by a time-dependent external force on the secondary structure level. Two recent unfolding RNA experiments carried out by Bustamante group are mainly investigated. We find that, in contrast to popular two-state assumption about the RNAs free energy landscape along the molecular extension, the molecules used in the experiments do not present apparent energy barriers. The strong cooperative folding and unfolding transitions of the RNAs observed in the experiments and in our simulations arise from the interaction of the molecules and the light trap. In addition, we also investigate the properties of Jarzynski's remarkable equality, whose experimental test has received considerable attention.

INTRODUCTION

RNA folding, which includes thermodynamic and dynamic folding, is one of the central problems in biophysics. Many experimental techniques have been applied in the RNA folding, such as x-ray crystallograph, NMR spectroscopy, etc. Recently, single-molecule manipulation techniques developed in the past decade provide a novel way to solve the important problem (1–3). Their common principle is to exert an external force on single molecules, and then to record the molecular end-to-end extensions (force-extension curves). Compared to the conventional methods that only yield time-averaged snapshots of RNA structures, current single-molecule technique is able to track RNA folding and unfolding trajectories on a single-molecule level in real time (1). These experiments may shed new light on the RNA folding problem. Until now, RNAs investigated by this technique include simple secondary structure units of RNA (1), complex tertiary structural ribozyme (2), and 1540-base long 16S ribosomal RNA with pseudoknots (3).

The single RNAs manipulation experiments address a challenging issue for theorists: whether or how can we make use of the known secondary structural RNA knowledge to explain or predict the phenomena observed in the experiments? Many theoretical efforts have been devoted to understanding this issue (4–8). However, these theories or models are too simple to be useful in real experiments; free energy data about RNA secondary structures obtained by thermodynamic experiments (9) were often neglected. Moreover, they just only studied equilibrium behaviors. The intriguing nonequilibrium phenomena were not investigated seriously. Computational simulation should be a good choice to overcome these shortcomings. But we noted that, compared to enormous simulations for force unfolding proteins (10–13), simulations for RNAs were few (3) although the biological importance of the latter is the same as the former. Recently we proposed a continuous time Monte Carlo (CTMC) method to simulate force unfolding single RNAs to fill this gap (14,15). We mainly focused on thermodynamic and kinetic behaviors of single RNAs under a constant mechanical force. Our simulation particularly calculated folding and unfolding rate constants at different forces, and the results agreed with the experiment very well. Unfortunately, the method cannot correctly account for the fact that in many real experiments, e.g., the mixed ensembles that we are concerned with here, the mechanical force is often time-dependent (1,16). In our initial simulations, a rough velocity estimation $v = \delta x/\delta t$ was used (15), where δx is the displacement of a light trap after a dwell time δt. The major problem of the approximation is that we cannot uniquely choose δx and δt except that the ratio between them is a constant. Hence, nonequilibrium behaviors simulated by the method may be problematic. In this paper, we will show that this flaw could be naturally solved by a time-dependent CTMC approach.

The organization of the article is as follows. In the next section, we simply review our previous force unfolding single RNAs model, and then show how to extend it to the time-dependent force cases using the time-dependent CTMC method. Because the thermodynamic quantities, such as the free energy differences are necessary in the discussion of nonequilibrium behaviors, an exact partition function method for the mixed ensemble with vanishing velocity of light trap (6) is reformulated. In the Results section, we first study the RNA unfolding thermodynamics with the simulation and the partition function method to confirm the correctness of the simulation. In the following part, we investigate the kinetics of RNAs, which includes the force-extension curves in far-from equilibrium, the free energy reconstruction from nonequilibrium data, and a comparison of three free energy estimators. Finally we give our conclusion.

MODEL AND METHOD

Force unfolding RNA model

Same with the experiments (1,16), our discussion is restricted to RNA secondary structures. The mixed ensemble is sketched in Fig. 1. The position of the center of the light trap is moved according to a time-dependent

Submitted August 18, 2005, and accepted for publication November 28, 2005.
Address reprint requests to Fei Liu, E-mail: liufei@tsinghua.edu.cn.

0006-3495/06/03/1895/08 $2.00
doi: 10.1529/biophysj.105.070540

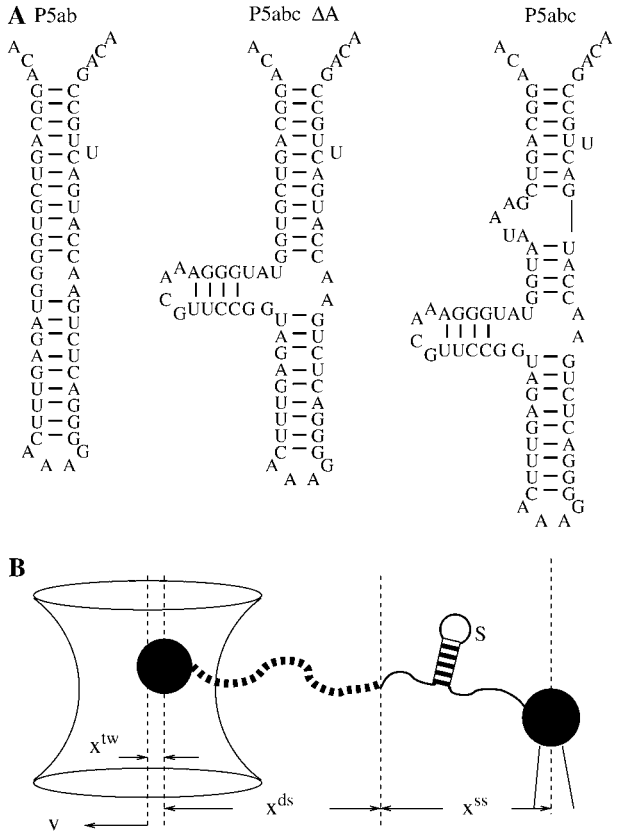

FIGURE 1 Sketch of the mixed ensemble and the native states of the RNA molecules P5ab, P5abcΔA, and P5abc studied in this work. The molecules are attached between the two beads (*larger black points*) with a RNA:DNA hybrid handle (*black dashed curves*). The center of the light trap is moved with velocity v. The total distance at time t is $z(t) = x^{tw} + x^{ds} + x^{ss}$.

relationship $z(t) = z_o + vt$, where $z(t)$ is the distance between the centers of the light trap and the bead held by the micropipette, z_o is the offset at time $t = 0$, and v is the constant velocity. We suppose that the changes of the extensions of RNA and the handle proceed along one direction, and the physical effect of the beads is negligible. Any state of the system at time t then can be specified with three independent quantities, the extension of the RNA x^{ss}, the end-to-end distance of the handle x^{ds}, and the RNA secondary structure S, i.e., the system in i-state at time $t(S_i, x_i^{ds}, x_i^{ss})_t$. Here x^{tw} is not included because the sum of individual extensions satisfies the constraint condition,

$$z(t) = x^{tw} + x^{ds} + x^{ss}. \tag{1}$$

Hence, unfolding of single RNA proceeds in space

$$S(\mathcal{S}) \times (0, l_{ds}) \times (0, l_{ss}), \tag{2}$$

where $S(\mathcal{S})$ is the set of all secondary structures of a given RNA sequence $\mathcal{S}$, l_{ds}, and l_{ss} are the contour lengths of the handle and the RNA molecule, respectively. To describe the RNA unfolding as a time-ordered series of the conformations in the space, a relation M, which specifies whether two conformations are accessible from each other by an elementary ''move'' must be reasonably defined. We proposed the following move set (15),

$$\begin{aligned}
&(S_i, x_i^{ds}, x_i^{ss})_t \rightarrow (S_j, x_i^{ds}, x_i^{ss})_{t'}, i \neq j \\
&(S_i, x_i^{ds}, x_i^{ss})_t \rightarrow (S_i, x_i^{ds} \mp \delta, x_i^{ss} \pm \delta)_{t'}, \\
&(S_i, x_i^{ds}, x_i^{ss})_t \rightarrow (S_i, x_i^{ds} \pm \delta, x_i^{ss})_{t'}.
\end{aligned} \tag{3}$$

The first kind of move is removal or insertion of single basepair while fixing the extensions x^{ds} and x^{ss}. The other two kinds are to, respectively, move the positions of the end of the handle and the end of single-stranded RNA with a small displacement δ, while the secondary structure is fixed simultaneously. Given the system state i at time t, the whole energy of it can be written as

$$E_i(t) = \Delta G_i^0 + u(x_i^{tw}) + W^{ds}(x_i^{ds}) + W^{ss}(x_i^{ss}, n_i), \tag{4}$$

where ΔG_i^0 is the free energy obtained from folding the RNA sequence into the secondary structure S_i, and the last three terms are the elastic energies of the optical trap, the handle, and the single-stranded part of the RNA, respectively (15). The light trap here is simply assumed to be a harmonic potential with a spring constant k_{tw}. The loading rate then is $k_{tw}v$.

Time-dependent CTMC algorithm

Given the conformational space, the RNA unfolding is modeled as a Markov process in it. Previous works (14,15,17) have demonstrated that CTMC simulation (18) was an excellent approach toward the stochastic process. As a variant of the standard Monte Carlo method, CTMC method is very efficient and fast because of lacking of waiting times due to rejection. Moreover, the ''time'' in the method could be real if the transition probabilities were calculated by first principles or empirically. Because we are considering time various external force, a time-dependent CTMC is essential (19).

The key formula in the time-dependent CTMC is that, given the system at i-state at current time t, the probability density $p(j, t'|i, t)$ that the next state is j and occurs at time t' is

$$p(j, t'|i, t) = k_{ij}^{t'} \exp\left(-\int_t^{t'} \sum_l k_{il}^{\tau} d\tau\right), \tag{5}$$

where k_{ik}^{τ} is transition probability from the i-state to the neighboring k-state at time τ (19), and the sum is over all neighbors of the current state. According to Eq. 5, the time t' for the next state to occur then can be obtained by solving equation,

$$r_1 = \exp\left(-\int_t^{t'} \sum_l k_{il}^{\tau} d\tau\right), \tag{6}$$

where r_1 is a uniform random number in the interval [0, 1]. For a time-independent situation, Eq. 6 reduces to the most common expression (15)

$$r_1 = \exp\left[-(t' - t)\sum_l k_{il}\right]. \tag{7}$$

The next state j is chosen if another uniform random number satisfies

$$r_2 \leq \sum_{l=1}^{j} k_{il}^{t'} \Big/ \sum_l k_{il}^{t'}. \tag{8}$$

For general k_{ik}^{τ}, solving Eq. 6 often requires time-consuming numerical integration and root finding. Here we simply assume that they have symmetric expressions, rule (20)

$$k_{ij}^{t} = \frac{1}{\tau_o} \exp\left\{-\frac{\beta}{2}[E_j(t) - E_i(t)]\right\}, \tag{9}$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, T is the temperature, and τ_o scales the time axis of the unfolding process. In addition to that, the transition probabilities obey the detailed balance condition locally in time (21,22), we find that the complicated formula of the right side of Eq. 6 has an analytical expression if the light trap is a harmonic potential. Great effort spent in the numerical integration therefore can be completely avoided.

Partition function method in equilibrium

If the moving velocity of the light trap vanishes, an exact partition function method can calculate the molecular average extension and the average force at a given distance z (6). The method is an extension of the partition function method proposed for RNA secondary structural prediction (23). Different from the experimental measurement of the free energy with slow pulling velocity (quasiequilibrium process) (16), we obtain the equilibrium information by this exact method. Considering coincidences of formulas and new physical quantities needed here, we rewrite the formulas proposed by Gerland et al. (6).

The central idea of the exact method is that the partition function over all secondary structures of a given RNA can be calculated by dynamic programming (23). Given the partition function $Q(i, j, m)$ on the sequence segment $[i, j]$ with exterior bases m, its recursion formula is given by,

$$Q(i,j,m) = \mathbf{1}\delta_{\mathrm{m,j-i+1}} + qb(i, \Delta + j - m) + \sum_{\mathrm{k=i}}^{\mathrm{j-1}} \sum_{\mathrm{l=1}}^{\mathrm{k-i+1}} Q(i,k,l) \times qb(k+1, l+\Delta+j-m), \quad (10)$$

where $\Delta = 2$, the partition function $qb(i, j)$ on the sequence segment $[i, j]$ for which the i and j bases are paired; Vienna Package 1.4 provides the calculation codes (24).

Let the partition function of the RNA molecule with n nucleotides at extension x (including the handle) be $Z_{\mathrm{n}}(x)$. Then the function can be written as

$$Z_{\mathrm{n}}(x) = \sum_{\mathrm{m=1}}^{\mathrm{n}} \int_0^{\mathrm{l_{ds}}} \int_0^{\mathrm{mb_{ss}}} dx^{\mathrm{ds}} dx^{\mathrm{ss}} \delta(x - x^{\mathrm{ds}} - x^{\mathrm{ss}}) Q(1,n,m) \exp[-\beta W(x^{\mathrm{ds}}, x^{\mathrm{ss}}, m)], \quad (11)$$

where

$$W(x^{\mathrm{ds}}, x^{\mathrm{ss}}, n) = W^{\mathrm{ds}}(x^{\mathrm{ds}}) + W^{\mathrm{ss}}(x^{\mathrm{ss}}, n). \quad (12)$$

The molecular free energy landscape along x then is

$$G_{\mathrm{o}}(x) = -k_{\mathrm{B}}T \ln Z_n(x). \quad (13)$$

To calculate the average force $\langle f \rangle$ and the average extension $\langle x \rangle$ at given distance z, we first have to calculate the free energy $G(z)$ of the whole partition function $\mathcal{Z}_{\mathrm{n}}(z)$ including the light trap by

$$\mathcal{Z}_{\mathrm{n}}(z) = \int_0^z dx Z_{\mathrm{n}}(x) \exp[-\beta u(z-x)], \quad (14)$$

and

$$G(z) = -k_{\mathrm{B}}T \ln \mathcal{Z}_{\mathrm{n}}(z). \quad (15)$$

Hence the $\langle f \rangle$ and $\langle x \rangle$ are, respectively, given by

$$\langle f \rangle = \partial G(z)/\partial z, \quad (16)$$

and $\langle x \rangle = z - \langle f \rangle / k_{\mathrm{tw}}$.

Parameters and measurement

We carry out simulations at the experimental temperature $T = 298\ K$ (1,16). The elastic parameters used here are: the persistence length of the handle $P_{\mathrm{ds}} = 53$ nm, $l_{\mathrm{ds}} = 320$ nm, $b_{\mathrm{ss}} = 0.56$ nm, Kuhn length of single-stranded part of RNAs $K_{\mathrm{ss}} = 1.5$ nm, and $k_{\mathrm{tw}} = 0.2$ pN/nm. We use the single-stranded DNA parameters for the single-stranded part of RNAs because they have similar chemical structures. The displacement $\delta = 1$ Å. The free energy parameters for the RNA secondary structures are from the Vienna Package 1.4 (24) in standard salt concentrations $[Na^+] = 1$ M and $[Mg^{2+}] = 0$ M. In addition to the standard Watson-Crick basepairs (AU and CG), GU basepair is allowed in our simulations. Formation of isolated basepairs is forbidden because of their instability. The instant force $f_{\mathrm{i}}(t)$ acting on the RNA molecule at i-state is calculated by

$$f_{\mathrm{i}}(t) = k_{\mathrm{tw}}[z(t) - x_{\mathrm{i}}^{\mathrm{ds}} - x_{\mathrm{i}}^{\mathrm{ss}}], \quad (17)$$

and the instant molecular extension is $x_{\mathrm{i}}^{\mathrm{ds}} + x_{\mathrm{i}}^{\mathrm{ss}}$.

RESULTS

Single RNAs thermodynamics

A comparison between our simulation in equilibrium and the partition function method is necessary to directly confirm correctness of our method. We simulate the average force-extension curves of the three RNA molecules in Fig. 1 with standard approach: average physical quantity A is calculated according to

$$\langle A \rangle = \tau^{-1} \int_0^{\tau} A(t) dt, \quad (18)$$

where $\tau = 10^6$; see the symbols in Fig. 2. We let $\tau_0 = 1$ for convenience here. The force-extension curves obtained by the exact method are plotted with different kinds of curves in the same figure. We can see that these two independent calculations agree very well.

Although the two methods agree with each other well, the values of the unfolding forces have apparent discrepancies with the experimental measurements. For example, in the absence of Mg^{2+} the values are 13.3, 11.3, and 8.0 pN for P5ab, P5abcΔA, and P5abc molecules in the mixed ensemble, respectively (1). It is not unexpected because we do not take account the effect of ionic concentration in our model. Hence, we choose a reasonable ionic correction of RNA free

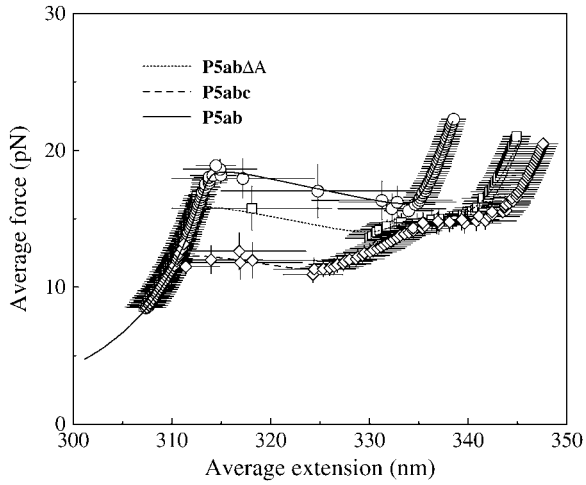

FIGURE 2 Comparison of the exact (*lines*) and the simulation (*symbols*) force-extension curves in equilibrium for P5ab, P5abcΔA, and P5abc in the mixed ensemble.

energies (25). Unfortunately, we still do not get good results; see Table 1. There are two possible causes leading to such discrepancies. One is that the ionic correction or free energy parameters for RNA are not precise enough to be used in the force unfolding cases. The other is that polymer elastic parameters are not good enough. We prefer the latter to the former. In addition to RNA free energy measured and tested for almost 40 years, the persistent length of small ssDNA in ionic environment is still debatable (28). For instance, we calculate the unfolding forces of the three molecules with $K_{ss} = 2.2$ nm and indeed find that they are closer to the experimental values. As a further demonstration, we also list other values measured in previous experiments and compare them with theoretical predictions in the same table.

Single RNAs kinetics

Force-extension curves

If one knows the force-extension curves in equilibrium, a way to check whether the unfolding proceeds in equilibrium or not is to see the coincidence of unfolding trajectories and equilibrium curves. Otherwise we have to check the coincidence of folding and unfolding trajectories. Fig. 3 *A* shows such an example: P5ab is stretched with the velocity $v = 5 \times 10^{-3}$ Å from the offset $z_o = 350$ nm to $z = 450$ nm, and then relaxed with the same velocity. Here the dimension of the velocity is distance instead of distance/time for the dimensionless $\tau_0 = 1$ assumed at the beginning. Apparently, the trajectories are not coincident, i.e., force-hysteresis, which indicates that the molecule is driven from thermodynamic equilibrium occurs.

Until now we did not point out what values of the velocity correspond to equilibrium or a near-equilibrium situation. It is necessary for in the real experiments, the equilibrium information, e.g., free energy differences, is obtained by a slow pulling velocity (16). To solve the problem, we unfold the three RNA molecules with two slower velocities, 1×10^{-4} and 1×10^{-5} Å. Because enormous data would be generated if the time trajectories were fully recorded, we only show the data per unit times 10^5 and 10^6 (see Fig. 3, *C* and *D*). For the faster velocity, we find that, except P5ab case, the unfolding forces for the others do not equal the equilibrium values, whereas for the slower case, the curves of simulations agree with the exact curves. It means that the unfolding of the three molecules with the later velocity has been in near equilibrium regime. In addition, two features in Fig. 3 *D* are of interest to us: 1), compared with the curves obtained by the time averaging in equilibrium, the curves recorded at time points are very rough even before and after the unfolding; and 2), although the whole extension $z(t)$ monotonically increases with time, the extensions of the molecules may still jump between two distinct values, such as P5ab and P5abc molecules. Indeed, similar phenomena were also observed in the experiment (1). They indicate the fluctuations of the extensions and RNA structures under the external force. We have known from the experiments (1,16) that P5abc unfoldings are near-equilibrium and far from equilibrium at the loading rates 2–5 and 34–52 pN/s, respectively (similar values for P5abcΔA). Our simulations showed that the unfoldings of the same molecule are, respectively, near-equilibrium and far from equilibrium at the loading rates $k_{tw}10^{-5}/\tau_0$ and $k_{tw}10^{-4}/\tau_0$. Let them equate to the experimental loading rates; correspondingly, we then estimate the constant $\tau_o \approx 10^{-7}$s. We will scale the time with this parameter in the following. The timescale obtained in this work is very different from previous $\tau_o \approx 10^{-5}$s (15), which was from the simulations in the constant force ensemble. A possible cause is that a different kinetic move set was used therein.

Fig. 4 shows 100 trajectories with two loading rates, 20 and 1000 pN/s for P5ab and P5abc molecules. The trajectories are stretched from the same offset $z_o = 350$ nm after thermal equilibrium until the terminal extension $z = 450$ nm. For both the loading rates, below and above the unfolding forces, the force-extension curves are dominated by the double-stranded handle. But the values of the unfolding forces apparently fluctuate and dependent on the rates and molecular types. When the pulling speed is faster, or the loading rate is larger (1000 pN/s), the average unfolding force increases correspondingly. This phenomenon has been theoretically predicted earlier (29). We note that at the same loading rate,

TABLE 1 The unfolding forces f_u of the different molecules under different experimental conditions

Molecule	Temperature (K)	Na^+ (mM)	Mg^{2+} (mM)	f_u^1(pN)	f_u^2(pN)	f_u^3(pN)	f_u^{exp}(pN)
P5abc	298	250	0	12.2	11.4	10.0	7.0–11.0
poly(dA-dU)	293	150	0	12.3	11.0	9.3	9.0
P5abcΔA	298	250	0	15.8	14.8	13.2	11.4 ± 0.5
P5abcΔA	298	250	10	–	15.4	13.8	12.7 ± 0.3
P5ab	298	250	0	18.4	17.4	15.7	13.3 ± 1.0
P5ab	298	250	10	–	18.0	16.2	14.5 ± 1.0
CG hairpin	293	150	0	25.8	24.4	22.4	17.0
poly(dC-dG)	293	150	0	25.1	23.8	21.7	20.0

The experimental data are from the published lectures (1,26,27). The theoretical values are from the exact partition function method; f_u^i, $i = 1, 2, 3$ represent the unfolding forces without the ionic correction, with the ionic correction on the free energy, and with the ionic and the Kuhn length corrections, respectively. We do not show the P5abc unfolding force for it is not reversible in Mg^{2+} due to the presence of tertiary interactions.

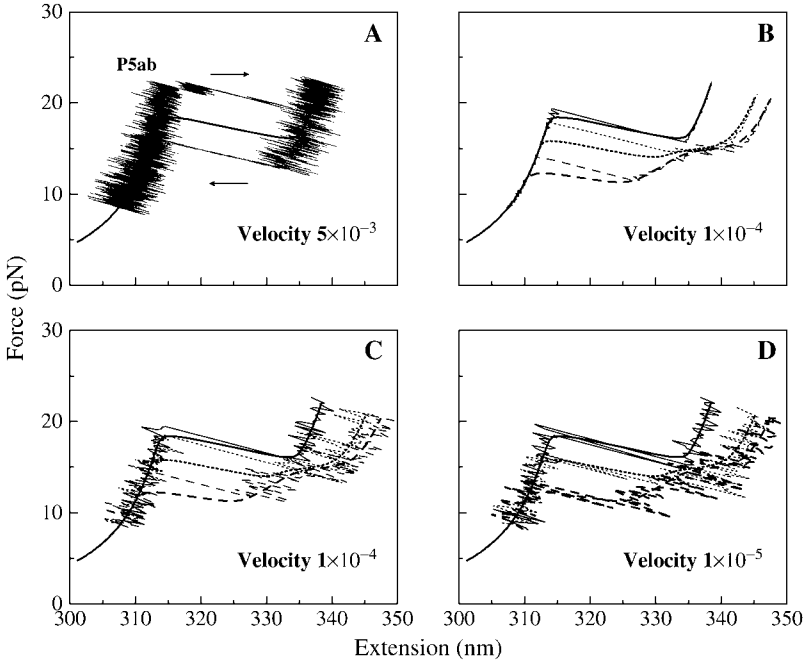

FIGURE 3 (*A*) One of the time trajectories of unfolding (*right arrow*) and refolding (*left arrow*) for P5ab with velocity 5×10^{-3} Å. Force hysteresis is observed. (*B*) The force-extension curves simulated by approximation velocity approach (15): $\delta x = 10$ Å, the dwell time $\delta t = 10^5$, hence the pulling velocity is 1×10^{-4} Å. (*C*) With the same velocity, the unfolding force-extension curves simulated by the time-dependent CTMC. We can see that the approximation approach works well if smaller dwell time chosen. The less noisy of the former is from a time-average in the dwell time (15). (*D*) The force-extension curves simulated by the current approach; the pulling velocity is 1×10^{-5} Å in this case. The smooth curves in these panels are the exact force-extension curves of the molecules in equilibrium.

the trajectories of P5ab are closer to its equilibrium force-extension curve than the trajectories of P5abc. It means that the relaxing process of the former is faster than the latter. This fact has also been observed in the experiment (1).

Free energy reconstruction

Recently, Hummer and Szabo (30) extended the remarkable Jarzynski equality (21) to extract unperturbed molecular free energy landscape $G_o(x)$ along the molecular extension x by the following expression

$$G_o(x) - G(t=0) = -\beta^{-1}\log\langle\delta(x - x(t))\exp(\Delta w_t)\rangle, \quad (19)$$

where $\Delta w_t = w_t - k_{tw}(x(t) - vt)^2/2$, $G(t=0)$ is the free energy of the whole system in equilibrium at initial time, and

$$w_t = k_{tw}v\left[vt^2/2 + z_o t - \int_o^t x(t')dt'\right]. \quad (20)$$

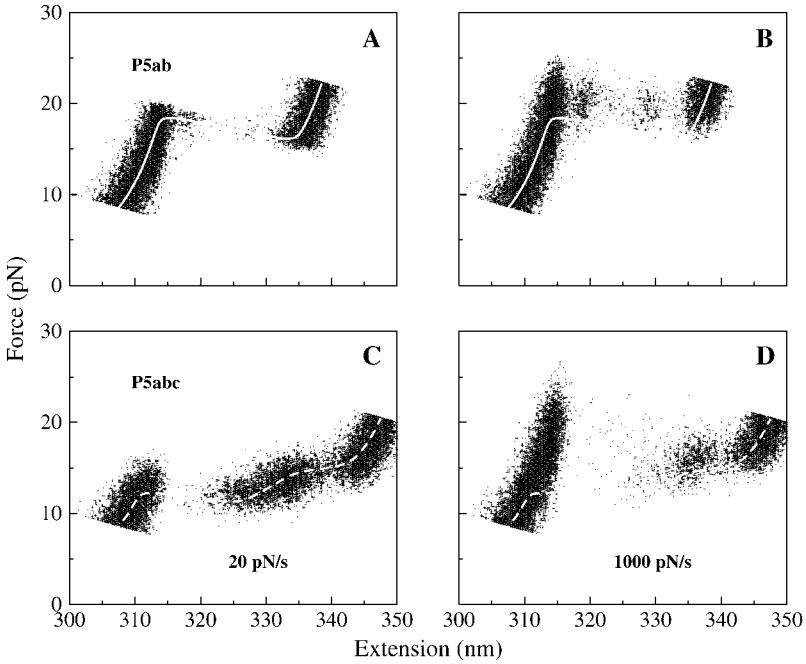

FIGURE 4 Unfolding trajectories of P5ab (*A*,*B*) and P5abc (*C*,*D*) with loading rates 20 and 1000 pN/s. Curves (superposition of 100 curves per figure) are represented by 100 points with the equal time interval. For clarity we do not connect them with lines. The white curves in these panels are the exact force-extension curves of the molecules in equilibrium.

It is interesting to construct the free energy landscapes of the real RNA molecules. Equation 19 implicates that the free energy landscape can be reconstructed just by one stretching. But considering that for finite stretching trajectories, we only sample a small window around the molecular equilibrium position at the whole extension $z(t)$. Therefore, a weighted histogram method was proposed (30),

$$G_o(x) - G(t=0) = -\beta^{-1}\log\frac{\sum_{t_i}\frac{\langle\delta(x-x_t)\exp(-\beta w_{t_i})\rangle}{\langle\exp(-\beta w_{t_i})\rangle}}{\sum_{t_i}\frac{\exp[-\beta u(x(t_i),t_i)]}{\langle\exp(-\beta w_{t_i})\rangle}}, \tag{21}$$

where the sum is over many time slices t', and the average is over the repeated trajectories at each given time slice. For each trajectory, we choose the discrete time $t_i = i\Delta t$, $i = 1, \cdots$, 100, here $\Delta t = 10/v$, i.e., the time moving the light trap 1 nm (or every point in Fig. 4).

Fig. 5 shows the finally reconstructed free energy landscapes for the two molecules at two loading rates 20 and 40 pN/s. The exact free energy landscapes obtained by the partition function method are also plotted there. We see that the precisions of the reconstructions are satisfactory. We also note that the landscapes are unexpectedly trivial: neither of them presents apparent energy barrier. Ritort et al. (31) have investigated Jarzynski's equality by modeling RNA molecules as a two-level system with an intermediate barrier. Our calculations contradict their assumption. The strong unfolding-refolding cooperativity observed in the experiments (1,16) and in our simulations actually arises from the interaction of the RNA molecules and the light trap; the addition of their potentials is a two-level system (see the respective *insets* in the figure). Therefore, the two-level system, although it is a good approximation in the RNA folding study, should not be simply copied to the force unfolding cases.

Free energy difference estimators

In addition to Jarzynski's equality (JE), there are two other common estimators of the free energy difference: the mean work (MW) and the fluctuation-dissipation theorem estimators (FD) (32). Similar to the experiment (16), a comparison among the three estimators using our simulation should be interesting. Instead of the molecular free energies, we will use the whole free energies (the molecules and the light trap) for simplicity. Their definitions are:

$$\text{JE estimator } \Delta G_{JE}(z) = -\beta^{-1}\log\langle\exp(-\beta w_z)\rangle_N,$$
$$\text{MW estimator } \Delta G_{MW}(z) = \langle w_z\rangle_N,$$
$$\text{FD estimator } \Delta G_{FD}(z) = \langle w_z\rangle_N - \beta\sigma_w^2,$$

where

$$\Delta G_i(z) = G_i(z) - G_i(z_o), \tag{22}$$

$i =$ MW, FD, and JE, and σ_w is the standard deviation of the work distribution (32) (we here replace time t by the whole extension z due to the linear relation between them). To get an intuitive observation about the estimators, we calculate the free energy differences between the estimators and the exact free energies,

$$\Delta G_i(z) - \Delta G(z). \tag{23}$$

The differences of P5ab and P5abc with the loading rates 20 and 40 pN/s, respectively, are shown in Fig. 6, *A* and *B*, and $N = 1000$.

There are two common features in the figure. First, the free energy differences for each estimator are not uniform along the molecular extension. For example for JE estimator, the differences are maximum around the unfolding extensions such as 415 nm for P5ab. We conclude that nonequilibrium behaviors of the same molecule are not uniform along its

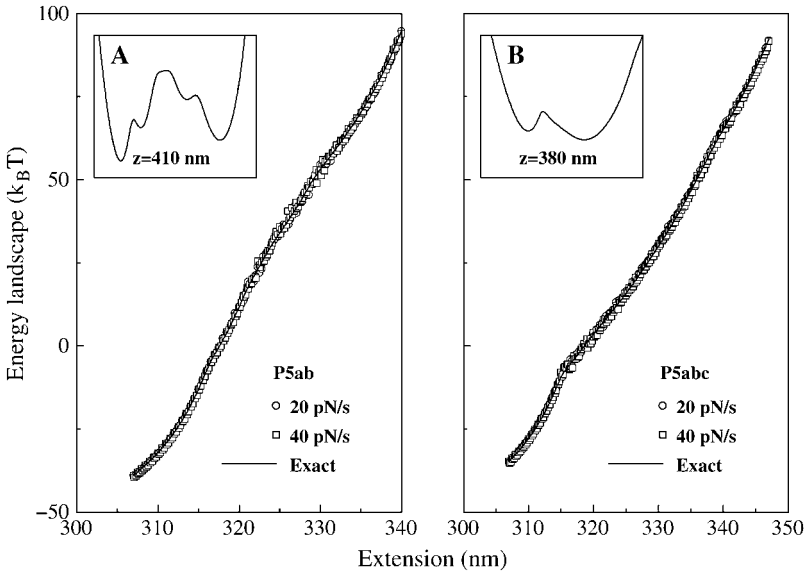

FIGURE 5 Comparison of the free energy landscapes of the two molecules P5ab and P5abc reconstructed from the Jarzynski equality and the exact landscapes calculated from the partition function method. The loading rates are 20 and 40 pN/s, and the number of trajectories for each case is 1000. The insets are the free energy landscapes of the whole systems composed of the RNA molecules and the light trap, which are from the partition function method. Note that we do not show the scales of the extensions and free energies.

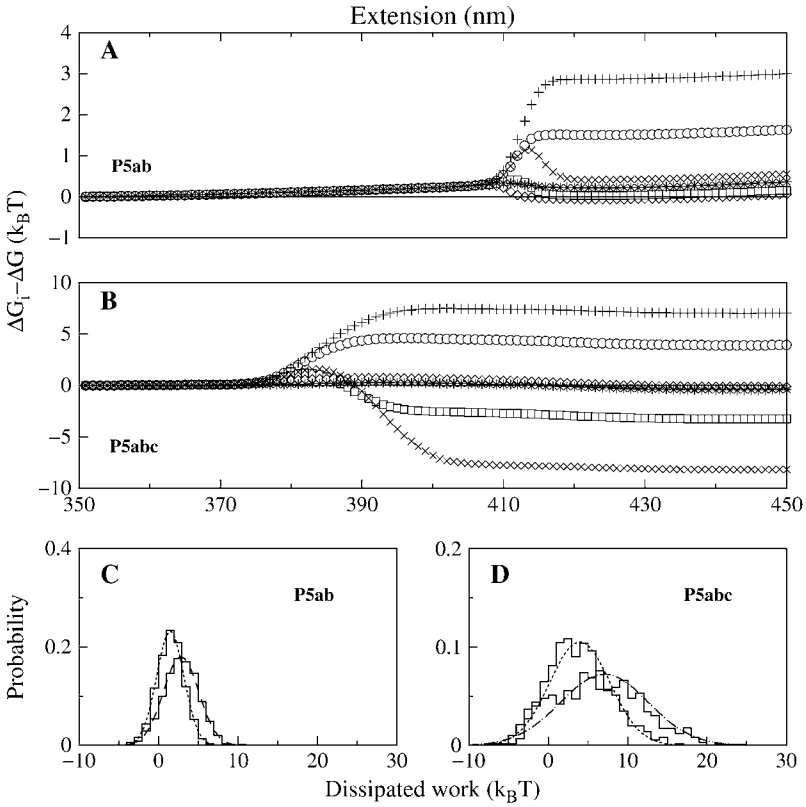

FIGURE 6 (*A*,*B*) The differences between the three free energy estimators and the exact energies for P5ab and P5abc at two loading rates 20 (*solid symbols*) and 40 pN/s (*crossed symbols*): MW (*circle and plus sign*), FD (*square and multiplication sign*), and JE (*diamond and star*); here $N = 1000$. (*C*, *D*) Histograms of the dissipated works at extension $z = 430$ nm for P5ab and P5abc molecules at the two loading rates. The solid lines are Gaussian functions with mean and variance from the same data set, the dotted lines are for 20 pN/s, and the dashed lines are for 40 pN/s.

extension, even if the RNA is unfolded with the same loading rate. Then for both the molecules, the JE estimator is always better than the MW at any loading rates. For the P5ab case, the FD estimator is more or less better than the JE at the extension $z > 415$ nm. This trend is more apparent as P5ab is unfolded with smaller loading rate 20 pN/s. In contrast, the JE estimator for P5abc is superior to the FD estimator over the entire extension range at the two loading rates.

These interesting observations can be understood from the distributions of the dissipated works

$$w_{\rm dis} = w_{\rm z} - \Delta G(z). \tag{24}$$

Fig. 6, *C* and *D*, show the histograms at a given distance $z =$ 430 nm, respectively. We find that for P5ab cases at the loading rates, the distributions agree well with Gaussian functions whose means and variances are obtained from the same data. This agreement is not unexpected: When a system is in the near-equilibrium regime, it always has a Gaussian dissipated work distribution, and in particular, an important equality, (32)

$$\sigma_{\rm w} = 2\beta^{-1}\langle w_{\rm dis}\rangle, \tag{25}$$

holds (in our data set $\sigma_{\rm w} = 2.86$ $(k_{\rm B}T)^2$ and $\langle w_{\rm dis}\rangle = 1.57$ $k_{\rm B}T$). Indeed, the force-extension curves of P5ab at lower rates have implicated this conclusion (see Fig. 3 *A*); whereas for the P5abc cases, we cannot obtain similar results because the system is driven far from equilibrium at the given loading rates (see Fig. 6 *D*).

Because the above analysis is carried out at a given *N*-value over a specific data set, we do not consider the errors caused by the value of *N* and the difference of samples. Gore et al. (33) have studied the issues in the large *N* limit and in the near-equilibrium regime. Our simulations also demonstrate correctness of their conclusions (F. Liu, unpublished data).

CONCLUSION

In this work, we improve our previous stochastic model to correctly take into account the time-dependent force by using the time-dependent CTMC method. We mainly consult with two recent RNA force unfolding experiments carried out by the Bustamante group (1,16). The simple RNA thermodynamic and kinetic properties under mechanical forces have been investigated. Combined with our previous effort for the constant force ensemble, our results show that, in contrast to protein cases, using the single polymer elastic theory and the RNA secondary structure free energy knowledge, we can successively simulate various behaviors of force unfolding RNAs under different experimental setups from equilibrium to far-from equilibrium. We hope that our simulations would be more useful in RNA unfolding studies in the future.

The computation of this work was performed on the HP-SC45 sigma-X parallel computer of the Institute of Theoretical Physics and the Interdisciplinary Center of Theoretical Studies, Chinese Academy of Sciences. F.L.

thanks Dr. F. Ye, R.-L. Dai, and Y. Zhang for their supporting in the computation.

F.L. was supported by the National Natural Science Foundation of China, grant No. 10447117.

REFERENCES

1. Liphardt, J., B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante. 2001. Reversible unfolding of single RNA molecules by mechanical force. *Science*. 292:733–737.
2. Onoa, B., S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, and C. Bustamante. 2003. Identifying kinetic barriers to mechanical unfolding of the T. thermophila ribozyme. *Science*. 299:1892–1895.
3. Harlepp, S., T. M. Marchal, J. Robert, J.-F. Leger, A. Xayaphoummine, H. Isambert, and D. Chatenay. 2003. Probing complex RNA structures by mechanical force. *Eur. Phys. J. E.* 12:605–613.
4. Montanari, A., and M. Mezard. 2001. Hairpin formation and elongation of biomolecules. *Phys. Rev. Lett.* 86:2178–2181.
5. Zhou, H., Y. Zhang, and Z. C. Ou-Yang. 2001. Stretch-induced hairpin-coil transitions in designed polynucleotide chains. *Phys. Rev. Lett.* 86:356–359.
6. Gerland, U., R. Bundschuh, and T. Hwa. 2003. Mechanically probing the folding pathway of single RNA molecules. *Biophys. J.* 84:2831–2840.
7. Lubensky, D. K., and D. R. Nelson. 2002. Single molecule statistics and the polynucleotide unzipping transition. *Phys. Rev. E.* 65: 0319171–0319177.
8. Liu, F., L. R. Dai, and Z. C. Ou-Yang. 2003. Theory for the force-stretched double-stranded chain molecule. *J. Chem. Phys.* 119: 8112–8123.
9. Mathews, D. H., J. Sabina, M. Zucker, and H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters provides robust prediction of RNA secondary structure. *J. Mol. Biol.* 288:911–940.
10. Rief, M., J. M. Fernandez, and M. E. Gaub. 1998. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* 81:4764–4767.
11. Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.
12. Klimov, D. K., and D. Thirumalai. 1999. Stretching single-domain proteins: phase diagram and kinetics of force-induced unfolding. *Proc. Natl. Acad. Sci. USA*. 96:6166–6172.
13. Socci, N. D., J. N. Onuchic, and P. G. Wolynes. 1999. Stretching lattice models of protein folding. *Proc. Natl. Acad. Sci. USA*. 96:2031–2035.
14. Liu, F., and Z. C. Ou-Yang. 2004. Unfolding single RNA molecules by mechanical force: a stochastic kinetic method. *Phys. Rev. E.* 70: 0409011–0409014.
15. Liu, F., and Z. C. Ou-Yang. 2005. Monte Carlo simulation for single RNA unfolding by force. *Biophys. J.* 88:76–84.
16. Liphardt, J., B. S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante. 2002. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*. 296:1832–1835.
17. Flamm, C., W. Fontana, I. Hofacker, and P. Schuster. 2000. RNA folding at elementary step resolution. *RNA*. 6:325–338.
18. Gillespie, D. T. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22:403–434.
19. Jansen, A. P. J. 1995. Monte Carlo simulations of chemical reactions on surface with time-dependent reaction-rate constants. *Comput. Phys. Commun.* 86:1–12.
20. Kawasaki, K. 1966. Diffusion constants near the critical point for time-dependent Ising models. *Phys. Rev.* 145:224–230.
21. Jarzynski, C. 1997. Equilibrium equality for free energy differences. *Phys. Rev. Lett.* 78:2690–2693.
22. Crooks, G. E. 1999. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E.* 60:2721–2726.
23. McCaskill, J. S. 1993. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*. 29: 1105–1119.
24. Hofacker, I. L. 2003. The Vienna RNA secondary structure server. *Nucleic Acids Res.* 31:3429–3431.
25. Cocco, S., J. F. Marko, and R. Monasson. 2003. Slow nucleic acid unzipping kinetics from sequence-defined barriers. *Eur. Phys. J. E.* 10:153–161.
26. Rief, M., H. Clausen-Schaumann, and H. E. Gaub. 1999. Sequence-dependent mechanics of single DNA molecules. *Nat. Struct. Biol.* 6: 346–349.
27. Bustamantle, C., S. B. Smith, J. Liphardt, and D. Smith. 2000. Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.* 10: 279–285.
28. Rouzina, I., and V. A. Bloomfield. 2001. Force-induced melting of the DNA double helix 1. Thermodynamic analysis. *Biophys. J.* 80: 882–893.
29. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.
30. Hummer, G., and A. Szabo. 2001. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA*. 98:3658–3661.
31. Ritort, F., C. Bustamante, and I. Tinoco, Jr. 1997. A two-state kinetic model for the unfolding of single molecules by mechanical force. *Proc. Natl. Acad. Sci. USA*. 99:13544–13548.
32. Hermans, J. 1991. Simple analysis of noise and hysteresis in (slow-growth) free energy simulations. *J. Phys. Chem.* 95:9029–9032.
33. Gore, J., F. Ritort, and C. Bustamante. 2003. Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. *Proc. Natl. Acad. Sci. USA*. 100:12564–12569.